Dyspraxia of speech in a British family: An acoustic study of diphthong production

Yan Kam Sum, Tom

A dissertation submitted in partial fulfillment of the requirements for the Bachelor of Science (Speech and Hearing Sciences), The University of Hong Kong, April 30, 2003.

Dyspraxia of speech in a British family: An acoustic study of diphthong production

Yan Kam Sum, Tom

Abstract

This paper investigated the acoustic properties of diphthongs produced by members of a family with a genetic speech and language disorder. The aim of the investigation was to provide acoustic and perceptual data on the diphthong and vowel errors produced by the affected members. Formant frequencies and vowel/diphthong durations were calculated from multiple repetitions of carrier phrases containing twenty British vowel/diphthong targets, including diphthongs in open and closed syllables, and tense and lax vowels in closed syllables. Utterances were collected from eight subjects (two unaffected and six affected members) from the second and third generations (19 to 58 years old). Perceptual analysis revealed diphthong errors of diphthongization, diphthong reduction and substitution, and tense rather than lax vowels were affected. Acoustic analysis indicated that the adult subjects were all able to contrast tenseness and laxness, and that compensatory lengthening of reduced diphthongs was observed. The reduction of a fronting diphthong was associated with a lack of or limited F2 upward movement, and the reduction of a backing diphthong was associated with a lack of or limited F2 downward movement. Acoustic analysis complemented perceptual analysis in defining and describing the various diphthong errors produced by the affected members.

INTRODUCTION

Data on the occurrence of developmental disorders of speech and language in families suggest that at least some form is familial, and hence inheritable (Lewis & Thompson, 1992; Tomblin, Records, Buckwalter, Zhang, Smith, & O'Brien, 1997). The study of genetically inheritable speech and language disorders has long provided researchers insights into the biological bases of language, its evolution and development. Though genetically based linguistic theoretical accounts of language acquisition or evolution in humans have been established for some decades (e.g. Chomsky's Language Acquisition Device, 1968), the definition of behavioral phenotypes from pathological studies is still of considerable debate. Against this background, the investigation of the three-generation KE family has been highly illustrative.

The KE family was first described by Hurst, Baraitser, Auger, Graham, & Norrell (1990). Half of the members of the first three generations are affected by a disorder of speech and language that renders them sometimes agrammatical and often unintelligible. Hurst et al. (1990) reported that the affected members were suffering from a "severe form of developmental verbal apraxia". Although in other genetically transmitted speech and language disorders more boys than girls are affected, males and females in the KE family are equally affected. Different accounts of the family history (e.g. Fletcher, 1996; Watkins, Dronkers & Vargha-Khadem, 2002) indicated that there is no doubt on the classification of the affected and unaffected status for each family member among researchers, the family members themselves, and clinicians and teachers who have known the family over a number of years. The disorder typically manifests itself in early childhood and persists throughout adulthood.

Linkage mapping and brain scanning studies conducted on the KE family have provided genotypic and neuroanatomical background for describing the speech and language disorder present in the affected members. Pedigree analysis has led Hurst et al. (1990) to conclude that the disorder in the KE family was transmitted as an autosomal dominant monogenic trait. A genetic linkage study concerning the family localized the abnormal gene (SPCH1) to a 5.6centiMorgan interval in the chromosomal band 7q31 (Fisher, Vargha-Khadem, Watkins, Monaco, & Pembry, 1998). In the same year, a brain-imaging study revealed that there were significant abnormalities in both cortical & subcortical motor-related areas of the frontal lobe of the affected members, which could account for their pronounced verbal dyspraxia (Vargha-Khadem, et al., 1998) [APA style format for > 6 authors] ~~Watkins, Price, Ashburner, Alcock, Connelly, Frackowiak, Friston, Pembrey, Mishkin, Gadian, & Passingham, 1998~~).

Behavioral studies seemed to yield results that lead to different theoretical interpretations concerning the underlying nature of the speech and language impairment in this family. Gopnik and colleagues (Gopnik, 1990; Gopnik & Crago, 1991) claimed that the affected members of this family had an inherited underlying linguistic deficit that is primarily grammatical in character, and particularly a deficit in the use of inflectional morphosyntactic rules (e.g. changing word endings to mark tense or number). They used the term "dysphasia" to refer to the language of the affected members. The linguistic account for the speech and language problem was further elaborated in Fee's (1995) phonological characterization of the KE family. By comparing the phonological profiles of the affected and unaffected members in the family, she concluded that the affected members acquired the phonological inventory of English following the normal developmental sequence, but at an extremely delayed rate. Consonants in syllable-final position and clusters at syllable level were susceptible to deletion or substitution errors. She concluded that these problems are consequent on their inability to learn language-specific phonological rules relating to complex syllable structures.

In contrast, Vargha-Khadem, Watkins, Alcock, Fletcher, & Passingham (1995) reported praxic and nonverbal cognitive deficits in addition to selective grammatical impairments potentially present in the affected family members. After comparing performances on intelligence tests, grammar tests, and reproduction of orofacial movement sequences, they found that the affected members showed intellectual, linguistic and orofacial praxic impairments. A more recent study conducted by Watkins, Dronkers & Vargha-Khadem (2002) showed that a test of non-word repetition successfully discriminated the affected and unaffected family members. All of the accumulated evidence point onto the fact that the phenotype expressed by this genetically inheritable speech and language disorder in the KE family seemed broader than it was originally conceived. In other words, cognitive, linguistic, neurophysiological and motor functions are impaired to different extents among the affected individuals.

Though the genetically inheritable speech and language disorder present in this family seemed unique, there exist striking similarities in the characteristics exhibited by the affected members and by individuals affected by "developmental apraxia of speech" (DAS). DAS has been most frequently described as an inability to program, or coordinate sequential speech movements in the absence of neuromuscular deficits (see, e.g. Davis, Jacks & McKenney, 2001; Davis, in press; Hall, 2000; Marquardt, Jacks, Stettler & Hruska, 2001; Smith, Marquardt, Cannito & Davis, 1994; Sussman, Marquardt & Doyle, 2000 etc.). According to Pollock and Hall (1991), the concomitant features of DAS include family history, persistence of speech problems, nonverbal oral apraxia, language disorder, "soft" neurological signs, word-finding difficulties and academic learning problems. The first four features were specifically described previously (Gopnik & Crago, 1991; Vargha-Khadem et al, 1995; Watkins et al, 2002), but the latter three were lacking documentation for the affected members of the KE family.

Diagnosis of DAS is typically by identification of a group of speech and language characteristics, including prosodic abnormalities, vowel errors, and segmental variability (Davis, Jakielski, & Marquardt, 1998). Phonetic descriptions of vowel misarticulations in the speech of children with DAS have been well documented. Morley (1959) stated that the diphthong errors often involved omission of the second part of a diphthong. Pollock & Hall (1991) reported that although the five children investigated in their study exhibited individual patterns of errors, trends such as difficulty to produce tense/lax vowel contrasts, diphthong reduction and backing were observed. A three-year longitudinal study conducted by Davis, Jacks & McKenney (2001) in a group of three children with DAS revealed that their vowel production was not accurate at any sampling points between 4;6 and 7;7 years. Their errors included neutralization and omission of vowels and diphthong errors. In comparison, English vowels are typically acquired by 36 months of age for developing children (Vihman, 1996).

Vowel errors in one affected member of the KE family, *A*., were described by Fletcher (1996). In a brief monologue, the subject systematically reduced 90% of the fronting diphthong targets in closed syllables (i.e. /eɪ/ and /aɪ/) into the onset of those diphthongs. For example, the words *name* (/neɪm/) and *right* (/raɪt/) were produced as [nem] and [rat] respectively. However, acoustic documentation of speech features of the affected members of the KE family is currently not available. An acoustic characterization of the affected family members might serve as an objective justification of the auditory-perceptual descriptions of DAS and of the speech of the affected KE family members. It would also provide understanding towards the phonological knowledge of the affected KE family members through examining subphonemic contrasts.

This research study had two main goals:

1) To study the formant characteristics (i.e. the first and second formants F1, F2) of diphthongs, elicited from structured speech tasks, produced by affected and unaffected members of the KE family. This would provide the first detailed acoustic description of the phonetic behavior of the family, and allow detailed comparison between affected and unaffected members of the family.
2) To provide evidence from acoustic data for perceptual labeling of "diphthong reduction" and "tenseness/laxness confusion" in the DAS speech as reported in Pollock & Hall (1991), and to provide an understanding of the phonological knowledge of the affected KE family members, by examining formant frequencies and vowel duration.

METHOD

Subjects

The subjects consisted of six affected members and two unaffected members of the KE family. The affected group consisted of two members from the second generation, and four members from the third generation, with age ranging from 19 to 58, mean 33.5 years. Three of them were male, who were referred to as AN, ST, and TO, and the rest were female, who were referred to as DN, TR, and VA. The unaffected group consisted of two males from the third generation, who were respectively 19 and 22 years of age, and they were referred to as BO and DE.

Stimuli

For the description of the formant characteristics in this study, five non-centralized and non-rhotic English diphthongs, /eɪ/, /aɪ/, /ɔɪ/, /əʊ/, and /aʊ/ were included. Four sets of monosyllabic words were constructed within either a consonant-vowel (CV, open syllable) or a consonant-vowel-consonant (CVC, closed syllable) context, utilizing these five diphthongs. The first and second sets included the five diphthongs placed into open and closed syllables respectively. There was indication in the data reported in Fletcher (1996) that diphthongs in closed syllables were more problematic for the subject *A*., while diphthongs in open syllables were produced accurately. The third and fourth sets included five lax vowels (/æ/, /e/, /ʊ/, /ɔ/, and /ɪ/), and five tense vowels, (/i:/, /ɔ:/, /u:/, /a:/, and /ɜ:/), placed into a closed syllable context. The singleton vowels were included in order to make formant and duration measurement comparisons with the diphthongs. There were 20 targets in total.

Each target word was embedded in a carrier sentence "Please say ____ to me", and each subject was instructed to produce each target six times, producing a total of 120 productions from each subject. There is evidence that the affected members are significantly poorer than the unaffected ones in repeating non-words (Watkins et al, 2001), and thus real words were used for all targets in this study to eliminate the potential effect of word unfamiliarity on the vowel and diphthong production. Because of the decision to use real words, it was not possible to control the consonant context for all targets. Contexts such as /l__d/, as well as /r__d/ and /w__d/ were thus included. Table 1 summarized all 20 targets.

Procedure

The 20 carrier sentences were produced by a native English speaker, and recorded on a compact disc (CD) in two successive randomized orders. There was a period of five-second silence between any two sentences so

Diphthong in open syllable (DOS)	Diphthong in closed syllable (DCS)	Lax vowels in closed syllable (LCS)	Tense vowels in closed syllable (TCS)
Lay /leɪ/	Laid /leɪd/	Lad /læd/	Lead /lɪːd/
Lie /laɪ/	Lied /laɪd/	Led /led/	Lard /la:d/
Roy /rɔɪ/	Lloyd /lɔɪd/	Lid /lɪd/	Lord /lɔ:d/
Low /ləʊ/	Load /ləʊd/	Rod /rɔd/	Word /wə:d/
Row /raʊ/	Loud /laʊd/	Wood /wʊd/	Rude /rʊ:d/

Table 1 Summary of all 20 targets with their phonetic transcriptions (British English)

that the subjects were asked to repeat after the speaker during the five-second pause. Each subject spent approximately 6 minutes to go through two random repetitions of each target, and he or she was required to go through the CD three times in order to produce six repetitions of each target. Due to file damage, two productions of the subject DN were lost and only 4 productions of each target word (a total of 80) could be collected.

The stimuli were delivered to the subjects through headphones. The recording was made in a quiet room using Kenwood KMC-D5 minidisk recorder with its microphone. The recording level of the minidisk was monitored so that the utterances were recorded at a sufficient level without clipping in the signal. A mouth-to-microphone distance of 10 cm was maintained. All productions were transcribed by the author using broad transcription for further comparisons with the acoustic data. The intra-rater reliability was 0.95. With who ?

Acoustic Analysis

The minidisk recordings obtained in the United Kingdom were transferred to two compact discs (CD-ROM), and were stored as ".wav" files on a *Microsoft* Windows 98 PC. The acoustic analysis was performed by using the PRAAT v4.0 software (Boersma & Weenink, 2001). A PRAAT script was used to segment the vowel/diphthong portion of the target stimuli and to estimate formant values. Formant estimates

Sound file	ANlord6								
Segment analyzed	Lord								
Segment duration	0.534								
Time steps	0%	12.50%	25%	37.50%	50%	62.50%	75%	87.50%	100%
F0	128.88	130.69	128.72	121.99	115.8	110.55	106.41	102.89	86.34
F1	454.48	424.07	380.05	365.75	336.5	296.07	314.09	254.61	238.99
F2	1024.21	901.89	779.59	800.41	901.16	1030.5	954.53	859.89	1007.96
F3	2082.77	2389.66	2518.4	2647.43	2441.01	2660.72	2819.2	2463.44	2286.8

Table 2 A sample table for duration measure and formant measures at nine time points

were calculated by firstly resampling the sample to a sampling rate of twice the value of Maximum formant frequency range (5000-8000Hz). After this, pre-emphasis was applied, and estimates of the first three formants were obtained through the LPC algorithm by Burg based on the routines "memcof" and "zroots" in Press, Teukolsky, Vetterling, & Flannery (1992). For a maximum frequency of 5000Hz, 10 coefficients were used by LPC algorithm; for a maximum frequency of 8000 Hz, 16 coefficients were used.

Each segment (i.e. the targeted vowels and diphthongs) was manually extracted by placing the vertical cursors at the onset and offset of the vocalic portion of each target. The onset of the segment was judged by visual inspection of the waveform and perceptual judgment of the preceding coarticulated consonant. The offset of the segment was judged by inspection of the waveform as the last voicing cycle. In cases of the continuous pulses stretching over the following coarticulated consonant, the offset was judged at the point where amplitudes of the audio waveform started to decrease. A sample of the output of the script for an utterance of the target "lord" is shown in Table 2.

Reliability measures

Ten percent of the target productions (i.e. 92) were randomly selected for reanalysis of segment duration, and the central (50% Time step) value of F1 and F2 as a measure of reliability by the author. Pearson's coefficient revealed high reliability for segment duration (r=0.98), and the central values of F1 (r=0.87) and F2 (r=0.86).

RESULT

Perceptual Analysis

All affected subjects produced vowel-related and consonant-related errors to varying degrees. Vowel-related errors included diphthong reduction, vowel/diphthong substitution and diphthongization, while consonant-related ones included gliding (production of /r/-initial targets as glide /w/), deletion and addition of final consonant, and initial consonant error.

For the purpose of this study, only the errors that pertained to vowels and diphthongs were discussed in greater detail. Table 3 summarizes the error patterns for all vowels and diphthongs, and the percentage of occurrence for all affected members. To illustrate, take the example of the target /laud/. A percentage of 1.00 for speaker VA means that she had consistently reduced /au/ to /a:/ in all utterances of this diphthong. In contrast, speaker AN reduced the DCS in only one of his six repetitions, thus giving a percentage of 0.17.

Segmental variability was frequently reported in the DAS literature (see, e.g. Davis, Jakielski, & Marquardt, 1998). For purpose of quantifying the degree of variability in the production of vowels, two measures of variability were computed, based on the formula designed in Marquardt, Jacks, Stettler, & Hruska (2001). Total Variance was given by the ratio of the total number of different vowel types to the total number of tokens, while Error Variance was obtained by dividing the number of different types of incorrect vowel/diphthong productions by the total number of incorrect productions. For example, the Total Variance for the two unaffected subjects BO and DE would be 0.167 (20/120), since all of the vowels and diphthongs

				Subject					
Target		R	EP	AN	DN	ST	TO	TR	VA
TCS	/lɔ:d/	/lɔɪd/	Diph	0.67		0.50			
		/lə:d/	Subs		0.25				
	/la:d/	/laɪd/	Diph	0.50					
	/wə:d/	/wɔ:d/	Subs				0.83		
DCS	/lɔɪd/	/lɔ:d/	Red	0.17					0.87
		/lə:d/	Subs+Red		0.25				
	/leɪd/	/led/	Red		0.25	0.17			
		/laɪd/	Subs				0.33		
	/laɪd/	/la:d/	Red						0.50
		/leɪd/	Subs	0.17	1.00				
		/neɪ/	Subs				0.33		
		/lʊɪ/	Subs				0.67		
	/ləʊd/	/laɪd/	Subs	0.17					
		/laʊ/	Subs	0.17					
		/lə:d/	Subs+Red		0.75				
		/lə:ʊd/	Subs		0.25				
		/lə:/	Red				0.67		
	/laʊd/	/la:d/	Red	0.17				0.33	1.00
DOS	/laɪ/	/leɪ/	Subs	0.17	0.75				
		/leɪd/	Subs		0.25				
		/lʊɪ/	Subs				1.00		
		/la:/	Red						0.17
	/raʊ/	/ra:/	Red						0.50
	/rɔɪ/	/weɪ/	Subs		0.75				
		/rɔ:/	Red						0.33
	/leɪ/	/laɪd/	Subs	0.17					
		/lʊɪ/	Subs				1.00		
	/ləʊ/	/lɔd/	Red	0.17					
		/nə:/	Red				1.00		
		/laɪd/	Subs	0.17					
		/lə:ʊd/	Subs		0.25				

Keys: R-Realization, EP-Error Pattern, TCS-Tense vowel in closed syllable, DOS-Diphthong in open syllable, DCS-Diphthong in closed syllable, Diph-Diphthongization, Subs-Substitution, Red-Diphthong Reduction.

Table 3 Realization, error patterns and percentage of occurrence for all diphthong and vowel errors produced by the affected subjects.

were produced correctly; since both of them produced no vowel or diphthong errors, the Error Variance could not be computed.

Table 4 shows that variability differed among the affected subjects. TR and ST had the highest token

	Subjects					
Measures	AN	DN	ST	TO	TR	VA
Total no. of tokens	120	80	120	120	120	120
Total no. of vowel errors	16	23	4	35	2	20
Token Accuracy	**0.875**	**0.763**	**0.967**	**0.708**	**0.983**	**0.833**
Total no. of different vowel types	31	27	22	24	21	25
Total Variance	**0.258**	**0.338**	**0.183**	**0.200**	**0.175**	**0.208**
Total no. of different types of incorrect vowels	11	10	2	8	1	6
Error Variance	**0.688**	**0.526**	**0.500**	**0.229**	**0.500**	**0.300**

Table 4 Summary of Token Accuracy, Total Variance and Error Variance for affected subjects.

accuracy and lowest total variance among the affected members. AN and DN had the highest total and error variances, and thus they were considered as the most variable speakers among the affected subjects. Though VA and TO had lowest token accuracy, they had smaller error variances in comparison to AN and DN, meaning that they were consistent in the errors they made. To conclude, an affected speaker with a low token accuracy did not necessarily have a higher error variance, though one with a high token accuracy was obviously associated with a lower total variance.

Table 5 shows that all affected subjects demonstrated the three error patterns, namely diphthongization, substitution and diphthong reduction to varying extents. AN demonstrated all of the three patterns, while DN, ST & TO had two. Diphthongization seemed to account for larger proportion of errors demonstrated by AN and ST, while DN, and TO committed more substitution errors. On the other hand, VA and TR committed only diphthong reduction errors. In general, all affected members had diphthong reductions, though substitution accounted for nearly half of diphthong errors (49 out of 100). Substitutions occurred more often in diphthongs than for tense vowels. Diphthong reduction affected diphthongs in closed syllable (DCS) to a greater extent than diphthongs in open syllable (DOS). Diphthongization was only

		Subjects						
Error pattern	Vowel type	AN	DN	ST	TO	TR	VA	Subtotal
Substitution	TCS		0.043		0.143			6
	DCS	0.187	0.391		0.229			20
	DOS	0.187	0.348		0.343			23
Diphthong reduction	DCS	0.125	0.217	0.250	0.114	1.000	0.700	28
	DOS	0.063			0.171		0.300	13
Diphthongization	TCS	0.438		0.750				10
Total number of diphthong errors		16	23	4	35	2	20	Grand total=100

Table 5 The proportion of different error patterns within each affected subject

	Tokens of errors					
Diphthong type	/ɔɪ/	/eɪ/	/aɪ/	/əʊ/	/aʊ/	Subtotal
DCS	7	2	3	7	9	28
DOS	2	0	1	7	3	13
Order	/əʊ/>/aʊ/>/ɔɪ/>/aɪ/>/eɪ/					

Table 6 Distribution of diphthong reduction among different types of diphthongs

observed in tense vowels (TCS), but not in lax vowels (LCS). Table 6 shows that all diphthong types were affected, but backing diphthongs (/əʊ/ and /aʊ/) were more challenging for the affected subjects than fronting diphthongs (/ɔɪ/, /eɪ/, and /aɪ/). The diphthong /eɪ/ was the easiest diphthong to produce for affected members.

Acoustic Analysis–Durational Measures

The vowel duration was measured to determine whether the affected subjects signaled contrast of tenseness and laxness and whether they produced a subphonemic contrast for reduced closed diphthongs. The duration of tense vowels was significantly longer than that of lax vowels ($p<0.05$) by Wilcoxon rank-sums test (W_s) for all subjects. This showed that both affected and unaffected subjects were able to signal contrast between tenseness and laxness from the perspective of durational measure, as opposed to the finding in Pollock & Hall (1991). Table 7 summarizes the mean duration and mean coefficient of variation of all target types for each subject. The mean coefficient of variation gives a measure of variability in the segment duration. It shows that the affected speakers were similar to the unaffected ones in the variability of the

	Subjects							
Target type	AN	DN	ST	TO	TR	VA	BO	DE
DCS	0.443	0.431	0.347	0.403	0.234	0.258	0.330	0.287
DOS	0.428	0.410	0.337	0.411	0.277	0.246	0.322	0.269
LCS	0.159	0.204	0.135	0.215	0.141	0.111	0.162	0.133
TCS	0.403	0.381	0.276	0.328	0.230	0.237	0.273	0.246
Coefficient of variation	0.164	0.201	0.250	0.174	0.133	0.117	0.121	0.153

Table 7 The mean duration and mean coefficient of variation of all target types for each subject

duration of all productions.

A point-to-point analysis was utilized to determine whether lengthening of reduced diphthongs, in both open and closed diphthongs, was evident for the affected subjects. The duration of a particular reduced diphthong production was compared with the mean duration of the tense vowel that was the same as the first vowel component of the reduced DCS. In the case of lengthening, a "Y" was assigned; otherwise, an "N" was assigned. To illustrate, the fifth production of /lɔɪd/ by AN was reduced, and since its duration (0.395s) was shorter than the mean duration of the corresponding tense vowel /ɔ:/ in /lɔ:d/ (0.478s), an N was assigned. There were a total of 41 diphthong reduction errors and there were only about half the number of errors (20) for which the duration of the reduced diphthongs was longer than that of tense vowel counterparts. Not all of the subjects lengthened reduced diphthongs consistently. Therefore, the results suggest that at least some of the affected subjects were able to lengthen reduced diphthongs, as a compensatory strategy to signal the presence of a second component of the diphthong.

Acoustic Analysis–Formant Measures

It is well known that in general /ɪ/ has low F1 and high F2, /a/ has high F1 and low F2, and /ʊ/ has low F1 and F2 (Peterson & Barney, 1952). Acoustic studies on the production of English diphthongs, whether by normal or abnormal speakers, are rare in the literature. Acoustic data on diphthongs of London English was

	Male				Female			
	F1	F2			F1	F2		
	Onset	Offset	Onset	Offset	Onset	Offset	Onset	Offset
/eɪ/	587	413	1945	2130	581	416	2241	2204
/aɪ/	734	439	1117	2058	822	359	1275	2591
/ɔɪ/	477	443	824	1924	428	334	879	2520
/əʊ/	537	379	1266	1024	545	380	1573	1267
/aʊ/	780	372	1368	1074	901	403	1538	1088

Table 8 Formant frequencies for London English diphthongs (Values in Hertz)

reported by Cruttenden (2001, p.99), and detailed in Table 8. The data shows that in general all English diphthongs in this table have decreasing F1; while the fronting diphthongs /eɪ/, /aɪ/ and /ɔɪ/ have rising F2, the backing diphthongs /əʊ/ and /aʊ/ have falling F2.

In the following analyses, two types of F2/F1 plots were used. All of the F2/F1 plots represented the changes of the F1 and F2 values, averaged across four productions of DN and across six productions from the rest of the subjects, from the third to the seventh sampling points (25% Time step to 75% Time step) for DCS, LCS and TCS, and from the third to the ninth sampling points (25% Time step to 100% Time step). The first included two F2/F1 plots for each subject depicting the formant changes of /a:/ and /ɔ:/, the only tense vowels that were considered diphthongized following the perceptual analysis. The second type included eight F2/F1 plots depicting formant changes of all diphthongs for each individual subject to allow across-subject investigations on substitution and reduction errors.

Diphthongization—There were three productions of TCS /la:d/ by AN considered as diphthongized as /laɪd/. From Figure 1, it can be seen that the two unaffected subjects BO and DE, as well as the affected subjects VA and TR showed decreasing F1 and F2 over time. While there was a slight increase in F1 for ST, all of the rest of subjects showed an increase in F2. Indeed, the F2 frequency of /a:/ by AN was high enough to cross the phonemic boundary so his intended production of /a:/ was perceived as /aɪ/. In contrast, the F2 of

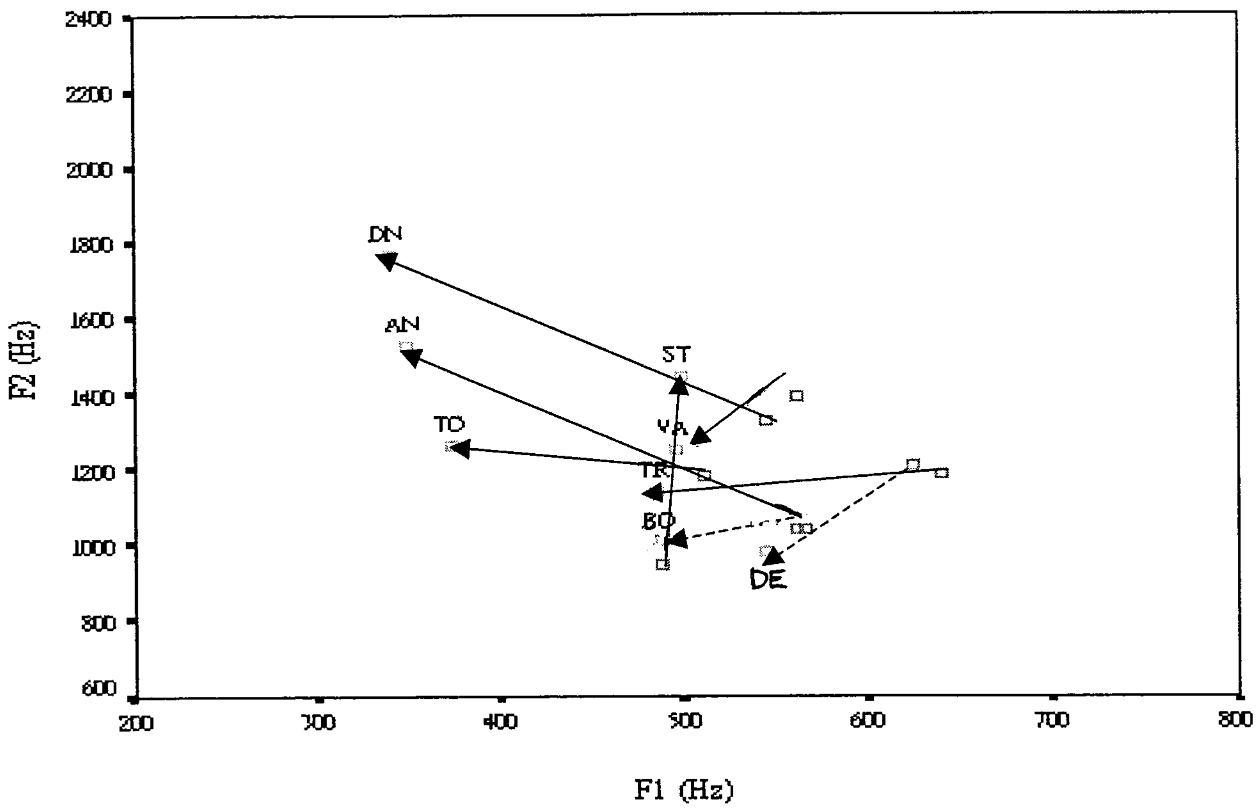

Fig. 1 Formant changes of tense vowel /a:/ for each subject

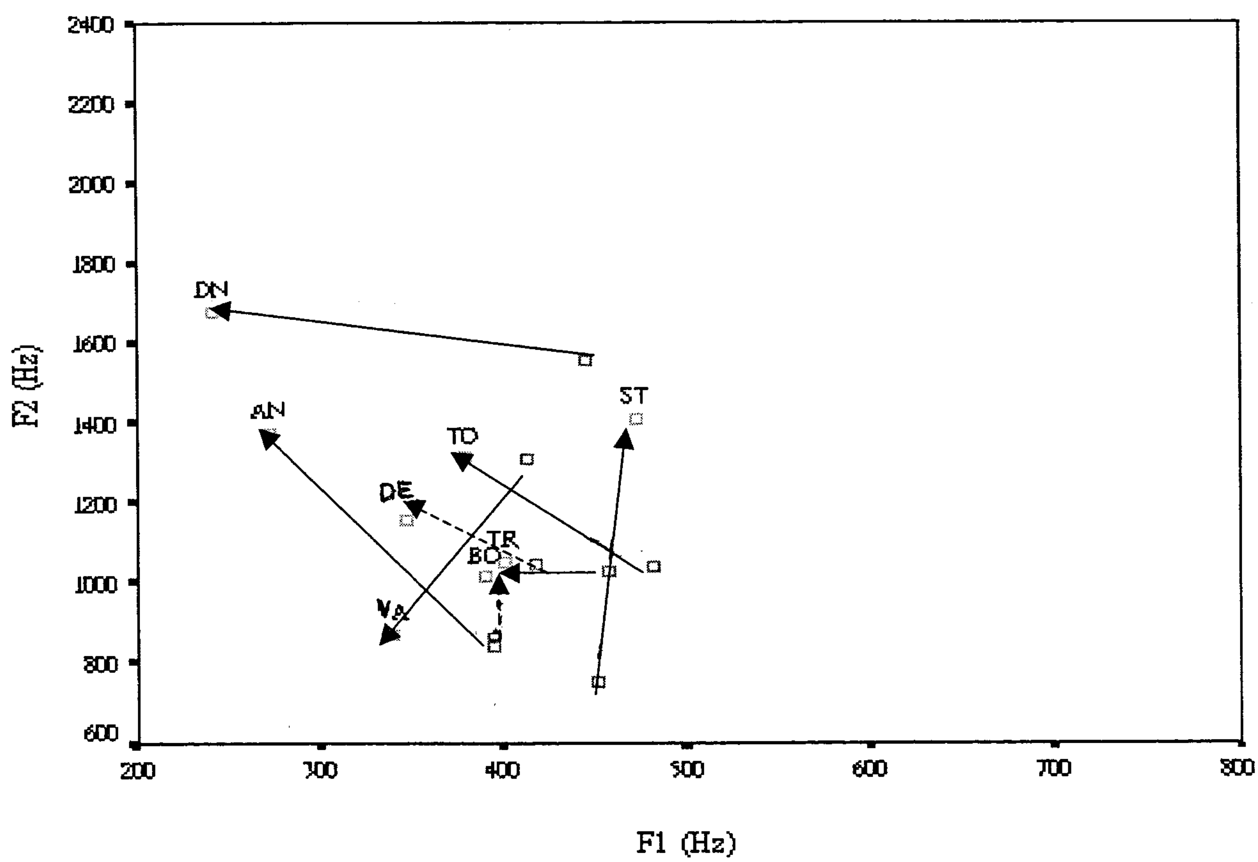

Fig. 2 Formant changes of the tense vowel /ɔ:/ for each subject

/a:/ by ST, TO and DE was not high enough to lead to the perception of diphthongization. Concerning the other tense vowel /ɔ:/, there were four productions of /lɔ:d/ by AN and three by ST considered as diphthongized to /lɔɪd/. From Figure 2 it can be seen that, with the exception of TR and VA, all speakers showed either a slight or a large increase in F2. For AN and ST, there was an increase of more than 500 Hz in the F2, which again, possibly led to their intended productions of /ɔ:/ being perceived as /ɔɪ/.

Diphthong reduction and Substitution—Figures 3 to 10 represented the formant changes of all diphthongs for each individual subject. The F1 and F2 changes of all diphthongs of the unaffected subjects BO, DE (see Figures 3 and 4) were similar to the normal data (c.f. Cruttenden (2001) in Table 8). The F1 and F2 changes of at least one diphthong for the rest of the subjects showed some qualitative differences, as discussed below.

The F1 and F2 changes of all diphthongs produced by ST and TR were similar to the normal data, with the exception that an increase in F2 of the /əʊ/ was observed (see Figures 5 and 6). This was expected for both of them since the only diphthong errors perceived in their productions included a few reduction errors (e.g. /leɪd/ realized as /led/, and /laʊd/ realized as /la:d/) and diphthongizations (i.e. /lɔ:d/ realized as /lɔɪd/). Though the trajectories of both /əʊ/ and /eɪ/ were similar for ST, no confusions were perceived in his productions because of overall lower F2 values for /əʊ/.

From the perceptual analysis, it was found that AN rarely produced diphthong reduction and substitution errors—only one out of six productions of each target was affected. This might explain the resemblance of the formant changes for all diphthongs to the normal data, as illustrated in Figure 7. The formant changes of the two backing diphthongs /əʊ/ and /aʊ/ were similar at starting point but diverged at the

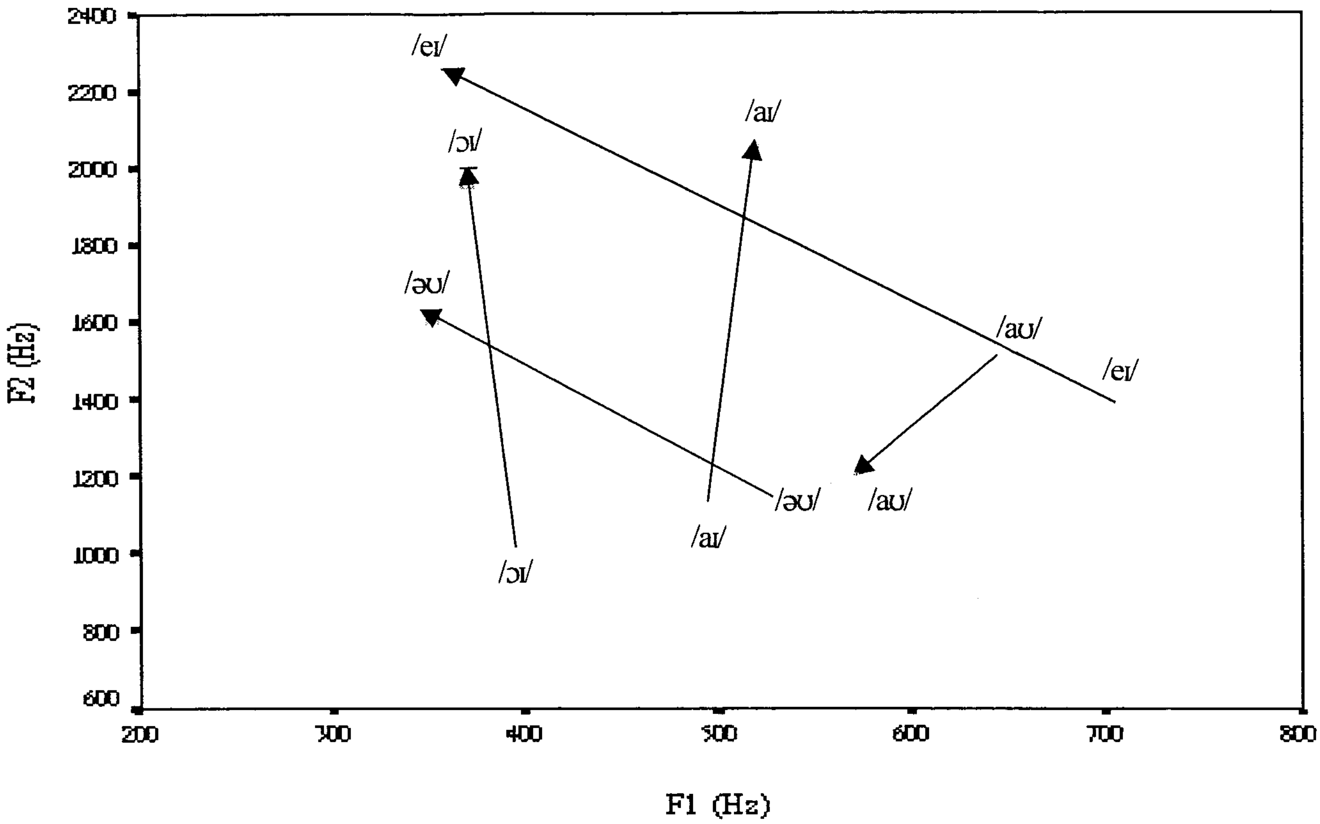

Fig. 3 Formant changes of all diphthongs for BO

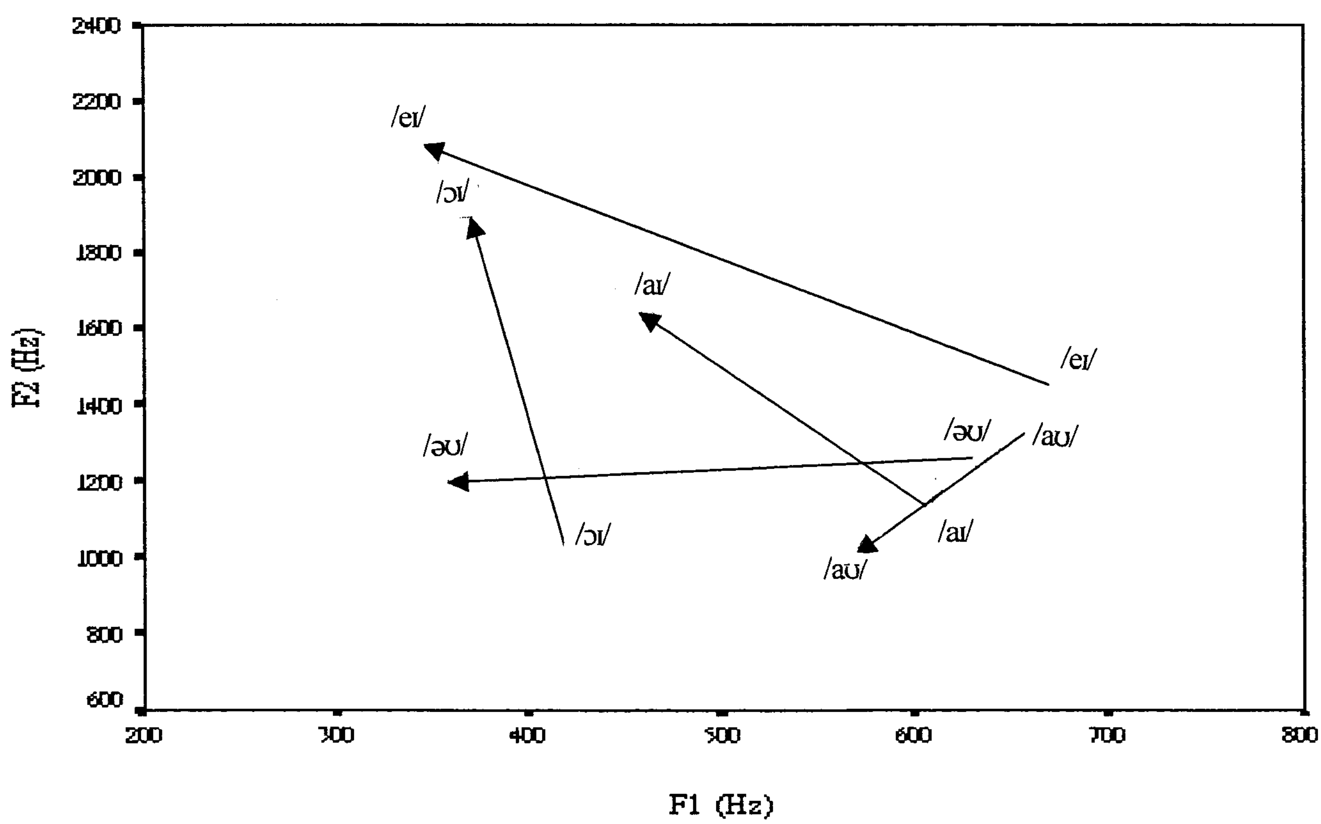

Fig. 4 Formant changes of all diphthongs for DE

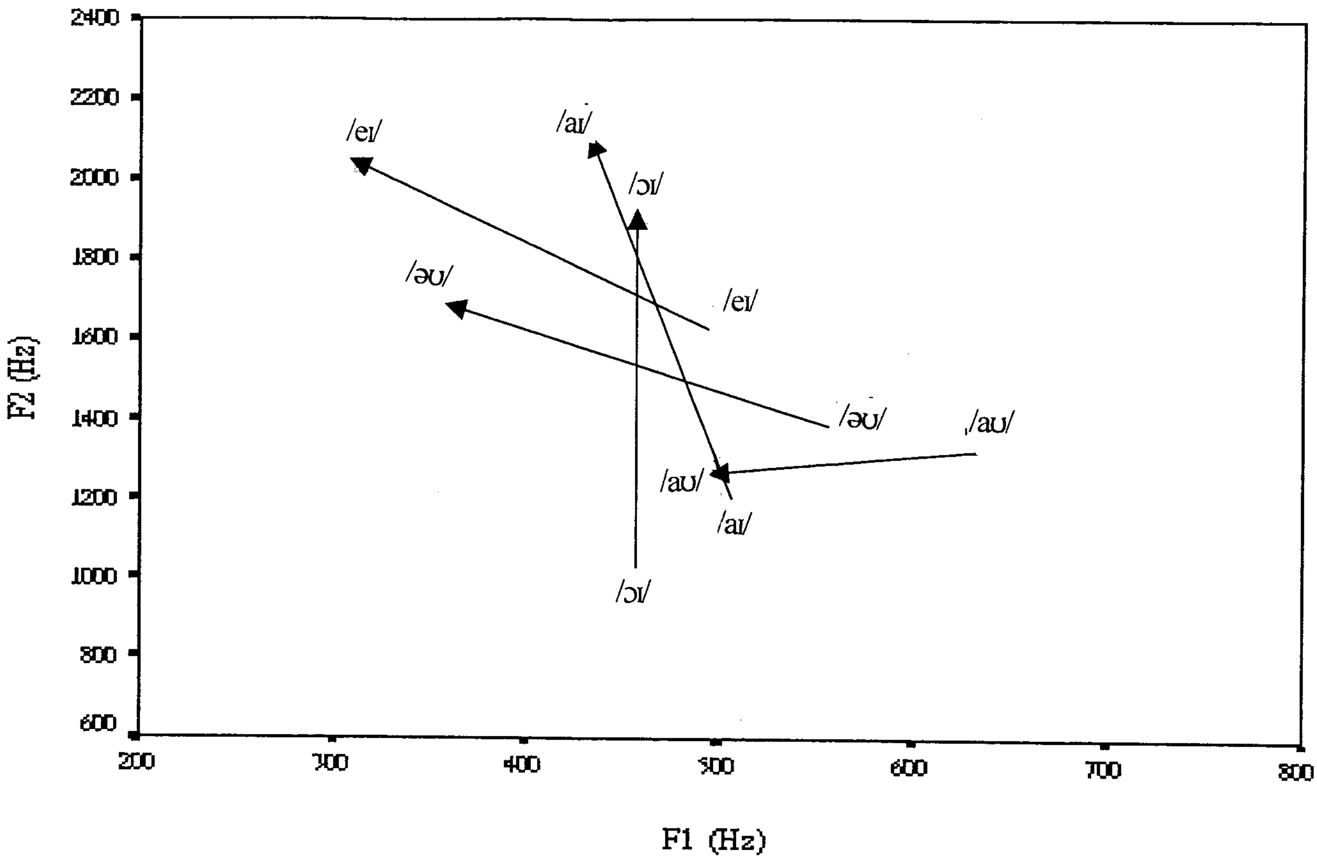

Fig. 5 Formant changes of all diphthongs for ST

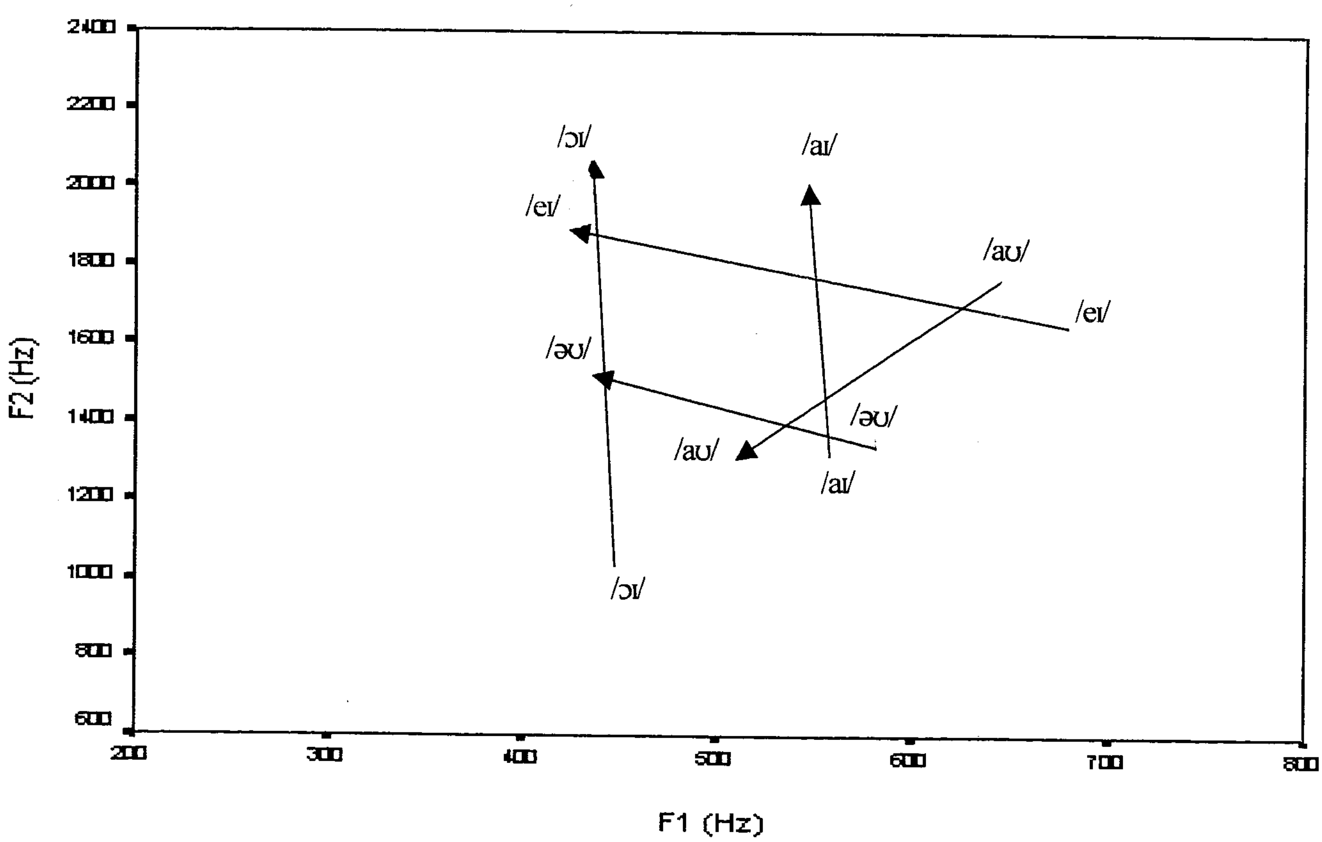

Fig. 6 Formant changes of all diphthongs for TR

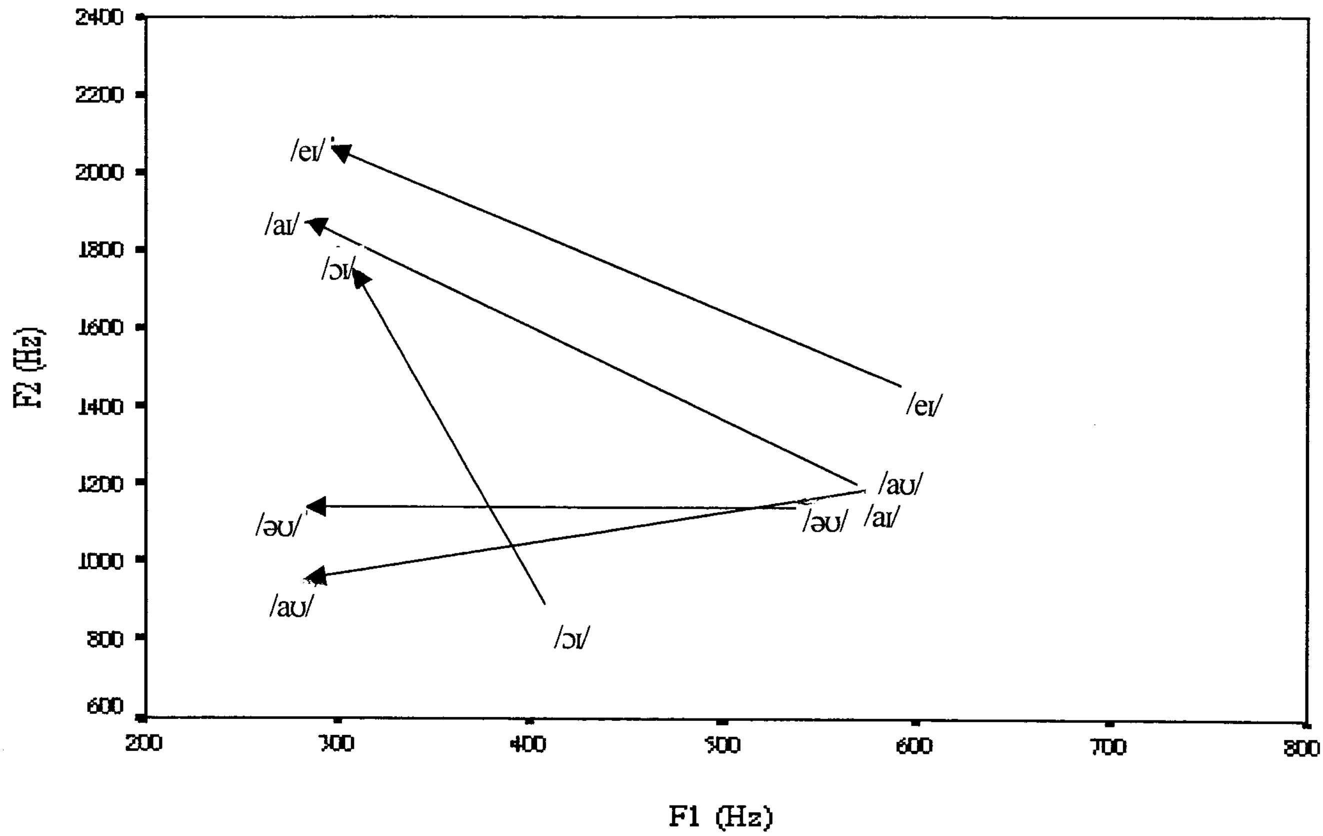

Fig. 7 Formant changes of all diphthongs for AN

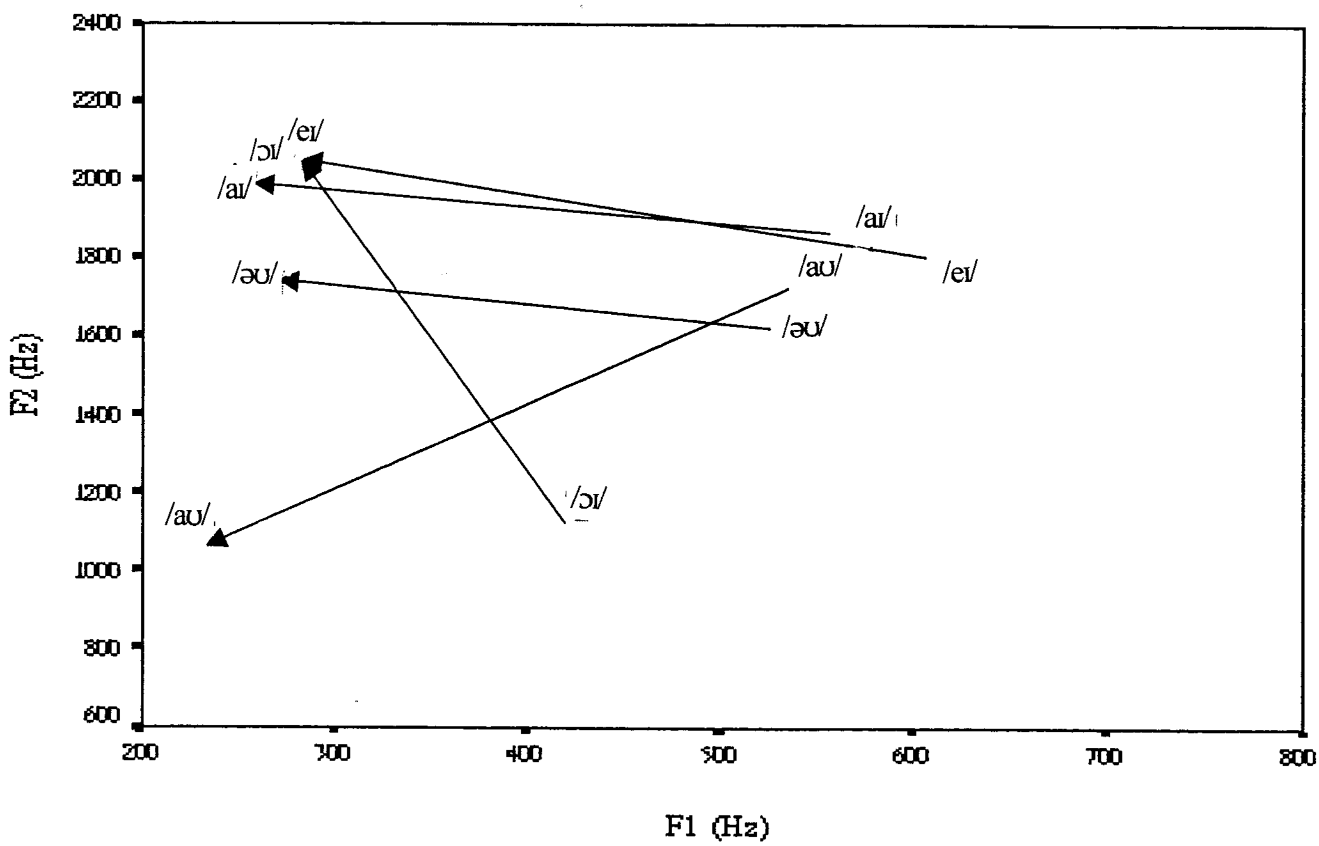

Fig. 8 Formant changes of all diphthongs for DN

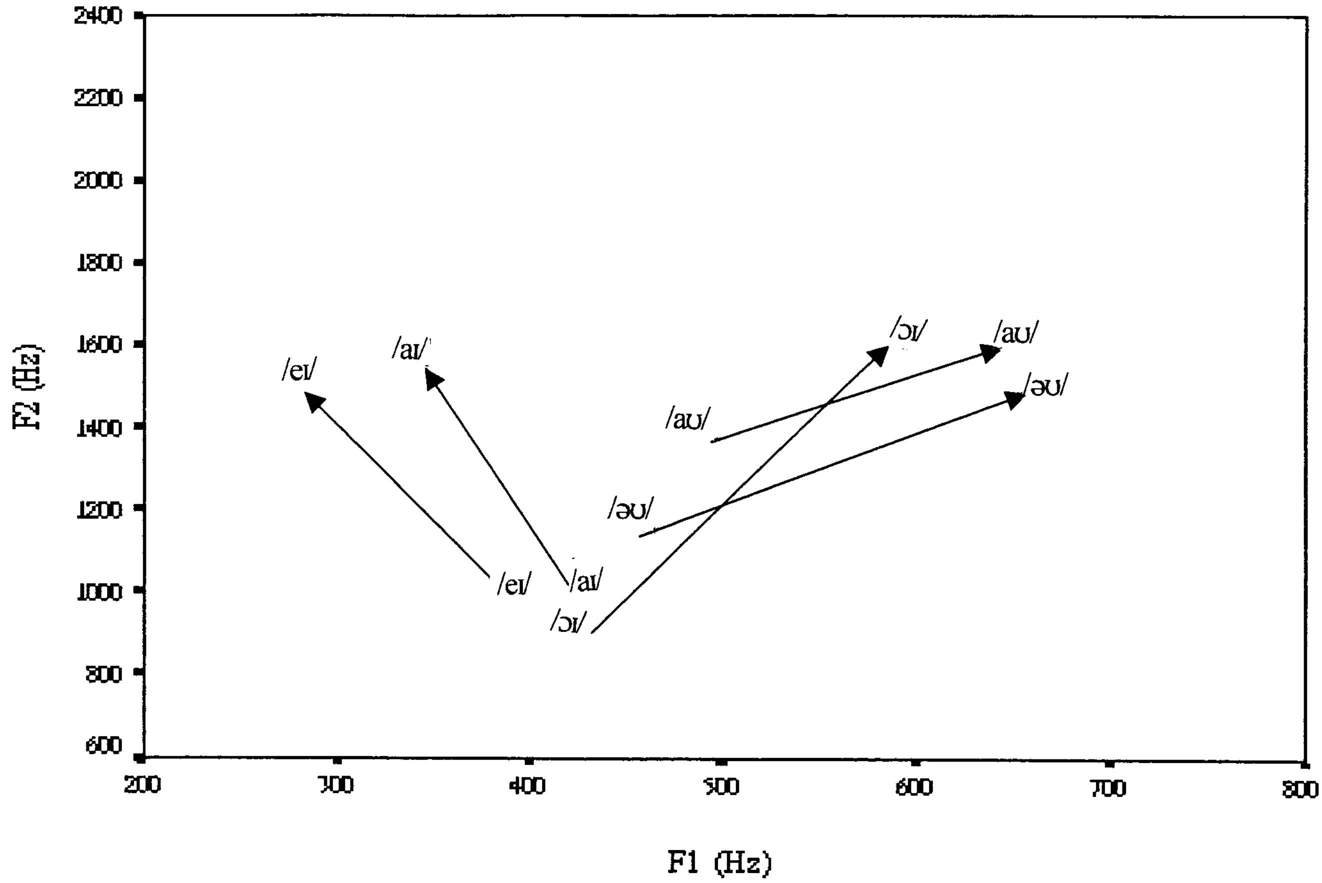

Fig. 9 Formant changes of all diphthongs for TO

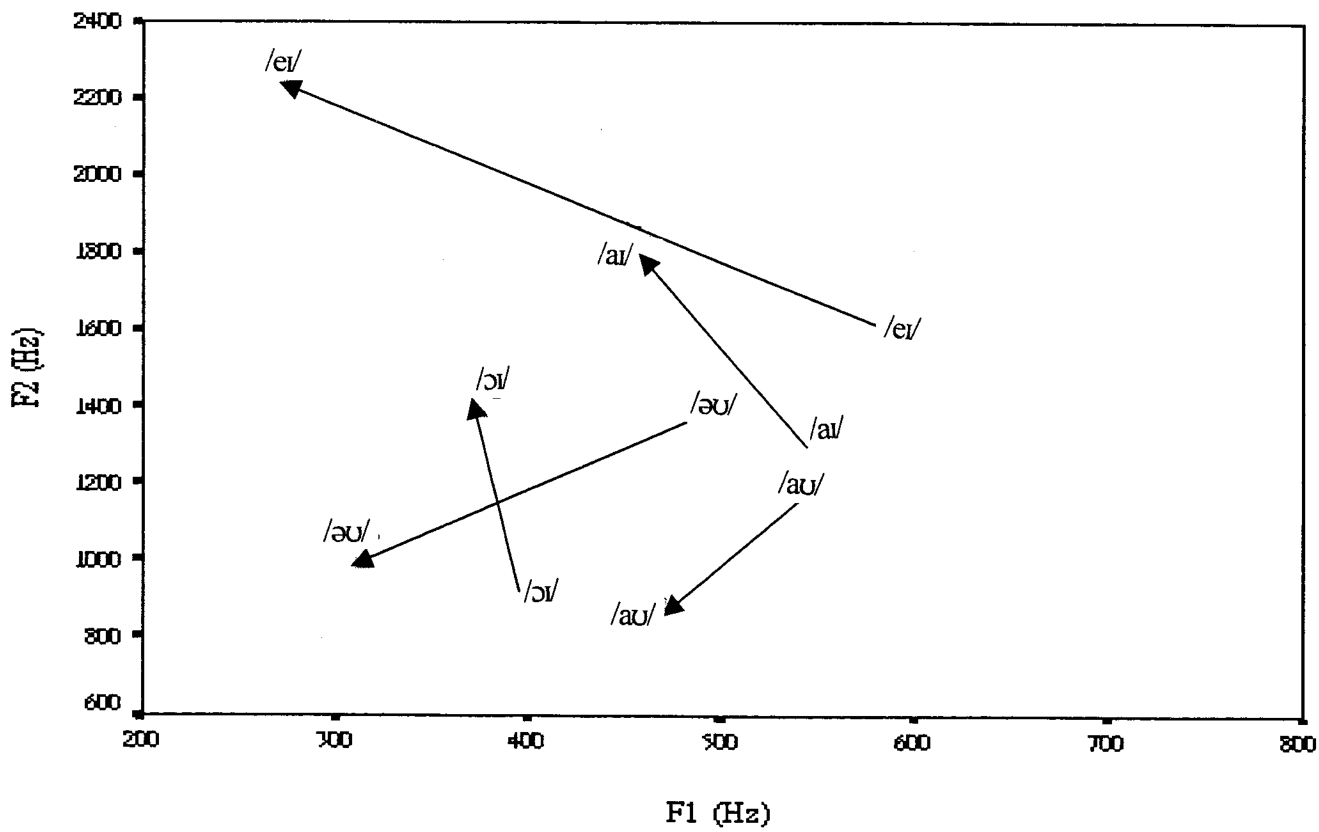

Fig. 10 Formant changes of all diphthongs for VA

end. This implicated confusion between the two diphthongs at the acoustic level that might not be realized at the perceptual level of analysis. In contrast, as could be seen from Figure 8, the closeness of the formant changes of /aɪ/ and /eɪ/ produced by DN was matched with the consistent substitution of /eɪ/ for /aɪ/ at perceptual level of analysis. The F2 change of /əʊ/ at the offset had a value similar to that of the central vowel /ə/ (i.e. a mid-high value of 1600-1800Hz for female), which might explain its perception as /ə/. Another prominent error observed in DN's production was the substitution of /eɪ/ for /ɔɪ/. However, the formant changes of /ɔɪ/ and /eɪ/ were not as close as those between /aɪ/ and /eɪ/. Therefore, in spite of difference in formant values, perception data showed confusion.

According to the perceptual analysis, TO was the most disordered speaker, having the smallest Token Accuracy. An interesting pattern was observed in the F1 and F2 changes for TO, as illustrated in Figure 9. The two fronting diphthongs /aɪ/ and /eɪ/ showed similar F1 downward and F2 upward movements, which might explain TO's confusion error between /eɪ/ and /aɪ/. The onset of low F1 and F2 values seemed to support the perception of substitution of /ʊɪ/ for both /aɪ/ and /eɪ/ in both open and closed syllable. At the same time, the backing diphthongs /əʊ/ and /aʊ/ and the third fronting diphthong /ɔɪ/ showed similar F1 and F2 upward movements. There was an atypical upward movement for /ɔɪ/ and a lack of a typical F2 downward movement for both /əʊ/ and /aʊ/. Therefore, results suggested confusion among these three diphthongs at acoustic level, despite that no substitution errors were identified at perceptual level of analysis.

Figure 10 displays the data for the subject that was most affected by diphthong reduction—VA. The F1 and F2 changes of all diphthongs of VA were also similar to normal data (c.f. Cruttenden (2001) in Table 8). At the perceptual level, VA demonstrated reduction of /ɔɪ/, /aɪ/ and was the only subject who reduced /aʊ/.

Though all of her transitions were similar to the normal data, the trajectories of /aɪ/, /oɪ/, and /aʊ/ seemed to be reduced in range in comparison with /eɪ/ and /əʊ/. In consideration of the three fronting diphthongs, the F2 value at the end point for /aɪ/ and /ɔɪ/ fell much below that of /eɪ/. On the other hand, a decrease in F1 and F2 of /aʊ/ was observed, which was typical of the transition of the backing diphthong /aʊ/, but the range of the decrease of F1 was relatively limited. The values of /aʊ/ showed that VA was still able to produce the /ʊ/ component subphonemically despite a perception of reduction.

Finally, the upward movement of F2 for backing diphthongs, despite no perception of diphthong reduction, was observed in /aʊ/ produced by TO and /əʊ/ produced by ST, TR, BO. There were two possible interpretations for this finding. These speakers produced either a subphonemic reduction of the /ʊ/ component, or a neutral component (i.e. /ə/) at the offset of the diphthong, thereby pushing up the F2 to a mid-high value (i.e. 1400-1600 Hz for male).

DISCUSSION

The data in this study provided an acoustic and a perceptual profile of the diphthongs produced by the affected members of a family affected by an inheritable speech and language disorder. All affected subjects produced vowel-related and consonant-related errors to varying degrees. Vowel-related errors included diphthong reduction, substitution and diphthongization, and consonant-related errors included gliding, deletion, addition of final consonants, and initial consonant errors. Variability in diphthong production was observed for all affected subjects. An affected speaker with a low token accuracy did not necessarily have a higher error variance, though one with a high token accuracy was obviously associated with a lower total variance Diphthong errors and variability in the production of vowels and diphthongs have been reported in previous studies on the vowel misarticulations of DAS children (Davis et al, 2001; Pollock & Hall, 1991; Smith et al, 1994). In this study, the affected members were similarly affected by diphthong reduction and substitution errors. This study demonstrated the pervasiveness of vowel disorders in adult affected members. Consonant errors observed in this study, gliding and final consonant deletion, were reported in Fee's (1995) phonological account of the KE family. However, final consonant addition observed in some of the productions of all affected members and the nasal realization of /l/ observed in TO have not been reported. It is too premature here to argue for or against any linguistically-based or phonetic account of these consonant errors, since only monosyllabic single words and a limited syllabic context were utilized in this study.

From an articulatory perspective, a diphthong requires a shift in vocal tract configuration during its production. In producing a diphthong in a closed syllable (DCS), the additional task of planning and, or executing an upcoming consonant appeared to overload the system to the extent that the diphthongal shift did not take place. This can explain the observation that diphthong reduction affected diphthongs in closed syllable (DCS) to a greater extent than diphthongs in open syllable (DOS). By the same token, it might be hypothesized that production of a diphthong, as opposed to a single vowel could have overloaded the production system, at least for some individuals, so that substitution affected diphthongs to a greater extent than tense vowels (e.g. as observed in the consistent production of /aɪ/ as /eɪ/ by DN). Again, from an articulatory perspective, the degree of mouth shape movement in production of /aɪ/ from one extreme quantal position to another (i.e. back to front) is greater than that in production of /eɪ/, where the mouth shape only moves from a close-mid position to a close position. This account is consistent with the observation that backing diphthongs seemed to be more challenging for all affected subjects than fronting diphthongs. The additional lip rounding in the production of the /ʊ/ component might have added burden to the system compared with the mere anterior tongue movement for the production of fronting diphthongs. However, this processing account might not well explain the diphthongization errors as observed in ST and AN, because production of a diphthong obviously involves a more complicated articulatory gesture than that of a tense vowel. In contrast, this might reflect a compensatory strategy with an attempt to signal the contrast of tenseness, since diphthongization was only observed in tense vowels (TCS), but not in lax vowels (LCS). The DAS children in the Pollock & Hall (1991) study also demonstrated difficulties in producing tense vowels. Physiological studies, such as palatographic analysis, might be useful to help investigate the articulatory movements involved in diphthongization. Such studies might also provide objective measures to explain the diphthong errors observed in this study, namely diphthong reduction and substitution, as well as to explore relationship between the vowel/diphthong errors and the previously reported praxic deficits in the KE family (Vargha-Khadem et al, 1995).

Durational and formant measures were investigated at the acoustic level of analysis. In Pollock & Hall's (1991) study, loss of contrast between tense and lax vowels was reported for DAS children. In the present study, the duration of tense vowels was significantly longer than that of lax vowels by the use of Wilcoxon rank-sums test (W_s). Therefore, both affected and unaffected subjects were able to signal contrast between tenseness and laxness from the perspective of durational measure, as opposed to the finding in Pollock & Hall (1991). Secondly, for about half of the diphthong reduction errors, the duration of the reduced diphthongs was longer than that of tense vowel counterparts. Therefore, the results demonstrated that at least some of the affected subjects were able to lengthen reduced diphthongs, as a compensatory strategy to signal the presence of a second component of the diphthong.

Formant measures illustrated the acoustic phenotypes for the diphthong errors produced by the affected members. Diphthongization of the tense vowels /a:/ and /ɔ:/ with the addition of a /ɪ/ component, as observed in ST and AN, was associated with an expected F2 upward movement. The reduction of a fronting diphthong was evidenced by a lack of or limited F2 upward movement, and the reduction of a backing diphthong seemed to be related to a lack of or limited F2 downward movement. Acoustic analysis complemented perceptual analysis in defining and describing the various diphthong errors produced by the affected members. For example, the realization of /laʊd/ as /la:d/ by VA was not supported by a lack of F2 downward movement, though the F1 and F2 changes in her production of /a:/ were similar to that of /aʊ/.

The results of this study showed that vowel and diphthong misarticulations remained a pervasive problem in the affected KE family members, as might possibly for other DAS grown-ups. English vowels are typically acquired by 36 months of age for developing children (Vihman, 1996), and vowels usually constitute smaller problem than consonants in speech delayed children. However, there was little agreement among researchers and clinicians concerning the role of vowel and diphthong errors in the differential diagnosis of DAS. For example, Shriberg, Aram & Kwiatkowski (1997) compared error consistency in children with suspected DAS to those with speech delay; they failed to find differences between the two groups. After further comparing prosody-voice profiles, they concluded that "the only linguistic domain that differentiates some children with suspected DAS from those with phonologically delayed is inappropriate stress" (p.313). Given the high unreliability in vowel transcription, acoustic studies constitute a promising approach for exploring diagnostic indicators in vowel disorders in other pathological groups, and acoustic phenotypes for genetically inheritable speech disorders.

Several questions remain unanswered by the present study. The effects of adjacent consonants, the number of syllables in the target word, the production of the affected subjects at conversational level were not investigated. Pollock and Keiser (1990) found that post-vocalic /l/ affected the accuracy of the preceding vowel for phonologically disordered children. The affected subjects in this study produced both consistent and inconsistent errors, and the effect of different consonantal context needed to be addressed. As reported by Pollock and Hall (1991), additionally, since another characteristic frequently reported for DAS is the increase in errors with increased utterance length or complexity, it would be interesting to investigate the effect of increasing syllable number within a word and the effect of eliciting the target at conversational level on vowel/diphthong accuracy and formant changes.

Acknowledgement

I am deeply grateful to my supervisors, Professor Paul Fletcher and Doctor Valter Ciocca, for their invaluable guidance and unconditional support. Thanks are also extended to the KE family members and the

XX University

colleagues of Developmental Cognitive Neuroscience Unit in London without whose participation the study would not be possible.

REFERENCES

Boersma, P., & Weenink, D. (2001). *Praat 4.0: A system for doing phonetics by computer*. University of Amsterdam: Phonetic Sciences Department.

Chomsky, N. (1968). *Language and mind.* New York: Harcourt, Brace & World.

Cruttenden, A. (2001). *Gimson's pronunciation of English.* London: Arnold.

Davis, B.L., Jakielski, K.J., & Marquardt, T.P. (1998). Developmental apraxia of speech: Determiners of differential diagnosis. *Clinical Linguistics & Phonetics, 12(1)*, 25-45.

Davis, B.L., Jacks, A., & McKenney, J. (2001). *Vowel production patterns in developmental apraxia of speech.* Paper presented at the 4th International Speech Motor Control Conference, Nijmegen, June.

Davis, B.L. (in press). Developmental apraxia of speech. In R. Kent (ed.), *Encyclopedia of communication disorders*. MIT Press.

Fee, E.J. (1995). The phonological system of a specifically language-impaired population. *Clinical Linguistics & Phonetics, 9(3)*, 189-209.

Fisher, S., Vargha-Khadem, F., Watkins, K.E., Monaco, A.P., & Pembry, A.E. (1998). Localization of a gene implicated in a severe speech and language disorder. *Nature Genetics, 18*, 168-170.

Fletcher, P. (1996). *Language impairment in a British family.* Paper presented at the Evolution of Human Language Conference, Edinburgh.

Gopnik, M. (1990). Feature-blind grammar and dysphasia. *Nature (London), 344*, 715.

Gopnik, M., & Crago, M. (1991). Familial aggregation of a developmental language disorder. *Cognition, 39,* 1-50.

Hall, P.K. (2000). A letter to the parent(s) of a child with developmental apraxia of speech: Part II The nature and causes of DAS. *Language, Speech, and Hearing Services in Schools, 31*, 173-175.

Holbrook, A., & Fairbanks, G. (1962). Diphthong formants and their movements. *Journal of Speech and Hearing Research, 5*, 38-58.

Lewis, B.A., & Thompson, L.A. (1992). A study of developmental speech and language disorders in twins. *Journal of Speech and Hearing Research, 35,* 1086-1094.

Marquardt, T., Jacks, A., Stettler, M., & Hruska, M. (2001). *Segmental variability in developmental apraxia of speech.* Paper presented at the 4th International Speech Motor Control Conference, Nijmegen, June.

at Logo... ?

Morley, M.W. (1959). Defects of articulation. *Folia Phoniatrica, 11*, 65-124.

Peterson, G.E., & Barney, H.L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America, 24*, 175-184.

Pollock, K.E., & Hall, P.K. (1991). An analysis of the vowel misarticulations of five children with developmental apraxia of speech. *Clinical Linguistics & Phonetics, 5(3)*, 207-224.

Pollock, K.E., & Keiser, N.J. (1990). Vowel systems of normally developing and phonologically disordered children. *Clinical Linguistics and Phonetics, 4*, 145-160.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., & Flannery, B.P. (1992). *Numerical Recipes in C: The art of scientific computing (2nd ed.)*. Cambridge: Cambridge University Press.

Shriberg, L.D., Aram, D.M., & Kwiatkowski, J. (1997). Developmental apraxia of speech: III. A subtype marked by inappropriate stress. *Journal of Speech, Language and Hearing Research, 40,* 313-337.

Smith, B., Marquardt, T.P., Cannito, M.P., & Davis, B.L. (1994). Vowel variability in developmental apraxia of speech. In J.A. Till, K.M. Yorkston, & D.R. Beukelman (Eds.), *Motor speech disorders: Advances in assessment and treatment (pp. 81-89)*. Baltimore, Maryland: Paul H. Brookes Publishing Co.

Sussman, H.M., Marquardt, T.P., & Doyle, J. (2000). An acoustic analysis of phonemic integrity and contrastiveness in developmental apraxia of speech. *Journal of Medical Speech-Language Pathology, 8 (4)*, 301-313.

Tomblin, J.B., Records, N.L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language and Hearing Research, 40*, 1245-1260.

Watkins, K.E., Dronkers, N.F., & Vargha-Khadem, F. (2002). Behavioral analysis of an inherited speech and language disorder: Comparison with acquired aphasia. *Brain, 125 (3)*, 452-464.

Vargha-Khadem, F., Watkins, K.E., Alcock, K.J., Fletcher, P., & Passingham, R.E. (1995). Praxis and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder. *Proceedings of the National Academy of Sciences of the United States of America, 92*, 930-933.

Vargha-Khadem, F., Watkins, K.E., Price, C.J., Ashburner, J., Alcock, K.J., Connelly, A., Frackowiak, R.S.J., Friston, K.J., Pembrey, M.E., Mishkin, M., Gadian, D.G., & Passingham, R.E. (1998). Neural basis of an inherited speech and language disorder. *Proceedings of the National Academy of Sciences of the United States of America, 95*, 12695-12700.

Vihman, M.M. (1996). *Phonological development: The origins of language in the child* Cambridge, MA: Blackwell Publishers.